Bayesian Analysis of Lung Cancer Mortality in the Presence of Misclassification

Mohsen Vahedi[1], Mohamad Amin Pourhoseingholi[2], Ahmadreza Baghestani[3], Alireza Abadi[3], Sara Sobhi[4], Zeinab Fazeli[5]

Abstract

Background: Lung cancer is an important cause of cancer mortality. Mortality is a familiar projection to address the burden of cancers. However, according to Iranian death registry, about 20% of death statistics were still recorded in misclassified categories. The aim of this study was to estimate lung cancer mortality in Iranian population, using Bayesian approach to revise this misclassification.

Methods: We analyzed National death Statistic reported by the Iranian Ministry of Health and Medical Education from 1995 to 2004. Lung cancer [ICD-10; C34] was expressed as the annual mortality rates/100,000 by sex and by age group. The Bayesian approach to correct and account for misclassification effects in Poisson count regression with a beta prior was employed to estimate the mortality rate of lung cancer in age and sex groups.

Results: According to the Bayesian analysis, there were between 20 to 30 percent underreported mortality records in deaths due to lung cancer, and its mortality rate increased through the recent years.

Conclusion: Our findings suggested a substantial undercount of Lung cancer mortality in Iranian population. Therefore, policy makers who determine research and treatment priorities on death rates should pay more attention to this underreported data.

Keywords: Lung cancer; Mortality; Bayesian analysis

Please cite this article as: Vahedi M, Pourhoseingholi MA, Baghestani AR, Abadi AR, Sobhi S, Fazeli Z. Bayesian Analysis of Lung Cancer Mortality in the Presence of Misclassification. Iran J Cancer Prev. 2013; 6(Suppl.):1-5.

1. Dept. of Epidemiology and Biostatics, Tehran University of Medical Sciences, Tehran, Iran
2. Gastroenterology and Liver diseases Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
3. Dept. of Biostatistics, Shahid Beheshti University of Medical Sciences, Tehran, Iran
4. Cancer Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
5. Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Corresponding Author: Mohamad Amin Pourhoseingholi, PhD;
Assistant Professor of Biostatistics
Tel: (+98) 21 22 43 25 15
Email: Amin_phg@yahoo.com
Received: 15 Oct. 2012
Accepted: 20 Dec. 2012
Iran J Cancer Prev 2013; 1(Suppl.):1-5

Introduction

Cancer is the third most common cause of death in Iran [1], and approximately 50,000 new cases of cancer occur annually in Iran [2].

Lung cancer is the most common fatal malignant disease worldwide [3] due to high incidence and poor prognosis. It is predicted to be the 7th cause of death and responsible for 3% of mortalities by 2030 [4]. Lung cancer is the fifth leading tumor in Iran. Although its incidence is low, it has been increasing steadily in both men and women during the recent years [5, 6].

Mortality is a familiar projection to address the burden of cancers. With regards to cancer mortality, data are important to monitor the effects of screening program, earlier diagnosis and other prognostic factors [7]. Cancer mortality data can be used to guide policy makers to setup cancer prevention programs. However, this aim needs reliable death registry systems that report complete death statistics annually with true trends. On the other hand, in developing countries, the analysis of death statistic subject to misclassification is a major problem in epidemiological analysis leading to biases estimates, and can therefore cause one to underestimate health risks [8]. The World Health Organization (WHO) has encouraged member states to introduce systems of death registration involving medical certification of cause of death.

Similar to other developing countries, Iranian mortality information is still incomplete [9]. According to Iranian death registry between 15% to 20% of death statistics were recorded in misclassified categories such as septicemia, senility without mention of psychosis symptoms and other ill-defined conditions, etc [10].

In statistical literature, two approaches are recommended for misclassification: 1) using a small validation sample [11]; 2) Bayesian analysis in which subjective prior information on at least some subset of the parameters are used to re-estimate death statistic [12, 13].

The aim of this study was to re-estimate lung cancer mortality rate for Iranian population, using Bayesian approach.

Materials and Methods

Data Sources:

The National Organization for Civil Registration (NOCR) and the Ministry of Health and Medical Education (MOH&ME) established death registration systems in Iran [9]. National death statistic reported by the MOH&ME from 1995 to 2000 (registered death statistics for Iranian population at the Information Technology and Statistic Management Center, MOH&ME), and from 2001 to 2004 (published by MOH&ME) [1, 10, 14] stratified by age group, sex, and cause of death (coded according to the 9th revision of the International Classification of Diseases [ICD-9]) were included in this analysis. Lung cancer [ICD-10; C34] was expressed as the annual mortality rates/100,000 by sex and age group (0-5, 5-14, 15-49 and >=50 years of age). Iran's population in 1995-2004 was estimated using the census from 1996 conducted by Statistics Centre of Iran and its estimation according to population growth rate for years before and after national census [15].

Statistical implementation:

The Bayesian approach is derived from models proposed by Stamey et al. to correct and account for misclassification in a Poisson regression [8]. Stamey's technique extended the model recently and proposed to overcome the problem of misclassification in cancer data [12, 13]. Pourhoseingholi et al. developed this technique to estimate mortality rate of colorectal cancer [16] and liver cancer [17]. We studied Iranian death statistic in a Bayesian Poisson regression using Stamey's approach to re-estimate mortality rate of lung cancer.

All analyses were carried out using S-plus. The misclassification probability estimate which is proposed in prior distribution was based on Iranian death registration which introduced 15% to 20% of misclassified records in total deaths. We assumed 20% misclassification (as a misclassification parameter, reported for death due to cancer without mentioned the exact name of the kind of cancer in Iranian death registry) with a beta prior to re-estimate death statistic of Lung cancer from misclassified groups.

Table 1. Mortality rate (per 100,000) for Lung Cancer stratified by sex group before and after adjusting for misclassification by Bayesian model

Year	Male		Female		Total	
	FR	BR	FR	BR	FR	BR
1995	1.53	1.87	0.64	0.70	1.10	1.35
1996	2.26	2.81	1.15	1.50	1.72	2.15
1997	2.79	3.38	1.30	1.57	2.07	2.48
1998	3.02	3.71	1.79	2.20	2.42	2.98
1999	4.16	4.90	1.69	1.99	2.96	3.58
2000	4.61	5.39	2.19	2.56	3.43	4.01
2001	5.06	6.16	2.46	2.93	3.79	4.62
2002	7.49	9.14	3.33	4.00	5.47	6.73
2003	11.5	13.57	5.46	6.44	8.60	10.23
2004	11.4	13.55	4.59	5.60	8.11	9.57

FR: Frequentist Rate, BR: Bayesian Rate

Table 2. Age specific rate (per 100,000) for Lung Cancer mortality before and after adjusting for misclassification by Bayesian model

Year	<5 Years		5-14 Years		15-49 Years		>=50 Years	
	FR	BR	FR	BR	FR	BR	FR	BR
1995	0.10	0.12	0.006	0.01	0.23	0.28	7.69	9.38
1996	0.18	0.22	0.03	0.04	0.44	0.53	11.6	14.27
1997	0.20	0.24	0.08	0.10	0.49	0.59	14.04	17.55
1998	0.25	0.29	0.05	0.06	0.44	0.53	17.12	20.89
1999	0.39	0.46	0.04	0.05	0.78	0.92	19.90	23.68
2000	0.44	0.51	0.06	0.07	0.95	1.15	24.08	29.14
2001	0.14	0.17	0.10	0.12	0.94	1.15	23.06	28.36
2002	0.00	0.00	0.05	0.06	1.31	1.55	36.75	44.10
2003	0.09	0.11	0.16	0.19	1.94	2.31	57.54	68.47
2004	0.26	0.31	0.11	0.13	2.00	2.44	54.71	66.75

FR: Frequentist Rate, BR: Bayesian Rate

Results

All death records due to Lung cancer from 1995 to 2004 were included in this study. The rate of Lung cancer mortality was classified by sex and age, generated from original database (Frequentist Rate), and their Bayesian corresponding projections (Bayesian Rate) appeared in Table 1 and Table 2. According to the Bayesian re-estimate, there were 20 to 30 percent underreported mortality records due to Lung cancer (Figure 1). The mortality rate of Lung cancer increased dramatically during these years (Figure1 and Table1); however, a slight decreasing was observed from 2003 to 2004. Moreover, Lung cancer mortality was higher among men (Table1 and Figure2), and mortality raised as age increased (Table2).

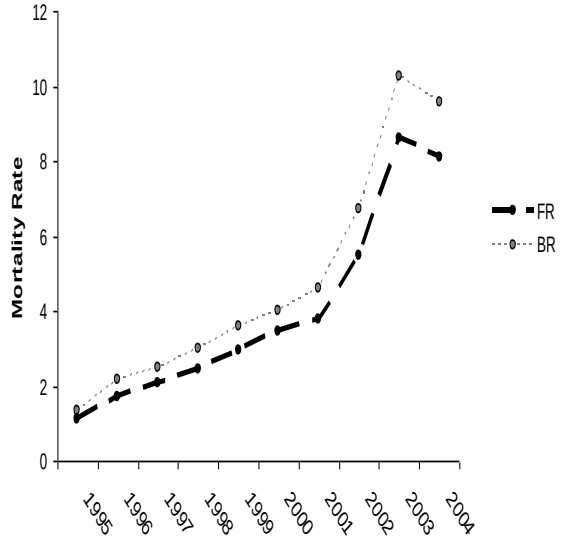

Figure 1. Trends of lung cancer mortality during the period 1995-2004
(FR: Frequentist Rate, BR: Bayesian Rate)

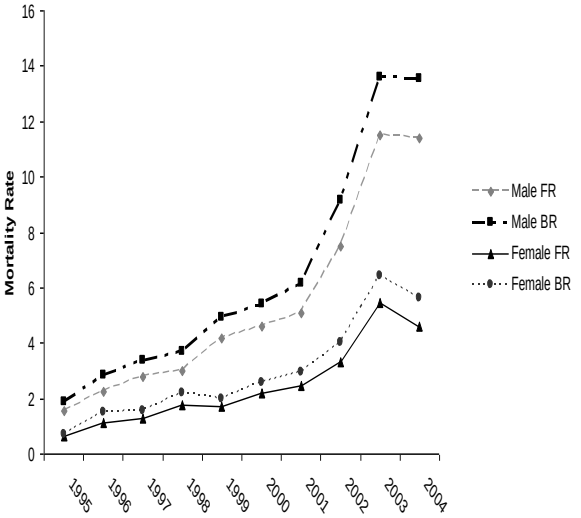

Figure 2. Trends of lung cancer mortality during the period 1995-2004 by sex groups (Frequentist Rate, BR: Bayesian Rate)

Discussion

Our results indicated that 20%-30% of lung cancer mortality remains underreported and suggested a substantial undercount of its mortality in Iranian population. Also, this study revealed that the trend of lung cancer mortality was dramatically increased in recent decade in Iranian population.

In the United States, age-adjusted lung cancer mortality rates have declined among males [18, 19]. Also in Canada, a decline in mortality rates has been observed [20]. In the United Kingdom, age-specific lung cancer mortality rates among men have declined since the 1980s. However, there was a rising rate for women [21].
Our results indicated a higher mortality in men compared to women, which could be due to more smoking among men [22, 23]. This is in contrast to developed countries which mortality of lung cancer is declining in men and rising in women [24, 25]. An Iranian study indicated a higher prevalence of lung cancer in males, older age and smokers with low and moderate socioeconomic classes [26].

Response misclassification of counted data for death statistics is still a problem in developing counties. In Iranian Death Registration System, data on causes of death are collected from various sources and have been assessed to be about 80% complete [9]. However, there is still up to 20% undefined death records that are categorized in wrong categories as misclassification.

Recently, Bayesian approach received much attention in the case of misclassification. Whittemore and Gong used this approach to estimate cervical cancer mortality rates [12], Sposto et al. developed this likelihood to assess the effect of diagnostic misclassification on non-cancer and cancer mortality dose–response in A-bomb survivors [13], Stamey et al. used Bayesian approach in data consisting of the number of deaths due to cancer and non-cancer among residents of Hiroshima and Nagasaki which were present during the atomic bombings in August of 1945 [8], and we used this technique to estimate the mortality rate of colorectal cancer [16] and liver cancer [17] according to Iranian death statistics.

Conclusion

This study provides comprehensive projection for burden of death due to lung cancer based on the national death registry, indicating that the mortality trend of this cancer was dramatically increased in the recent decade and there is a substantial undercount of lung cancer mortality according to the Bayesian model. Therefore, healthcare policy makers who determine research and treatment priorities on death rates as an indicator of burden of disease should pay special attention to this underreported data and set strategies against this cancer.

Acknowledgment

This study was sponsored by Gastroenterology and Liver diseases Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Conflict of Interest

The authors have no conflict of interest in this study.

Authors' Contribution

Mohamad Amin Pourhoseingholi conceived and designed this study; Mohsen Vahedi interpreted the results and drafted the manuscript; Ahmadreza Baghestani , Alireza Abadi and Sara Sobhi participated in writing and revising the manuscript; Alireza Abadi and Zeinab Fazeli contributed to data gathering and approved the final manuscript. All authors read and improved the final manuscript.

References

1. Naghavi M. Death report from 23 provinces in Iran. 1st edition. Tehran: Ministry of Health and Medical Education; 2004.
2. Mohebbi M, Mahmoodi M, Wolfe R, Nourijelyani K, Mohammad K, Zeraati H, et al. Geographical spread of gastrointestinal tract cancer incidence in the Caspian Sea region of Iran: spatial analysis of cancer registry data. BMC Cancer. 2008; 8:137.
3. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. CA Cancer J Clin. 2007; 57 (1): 43-66.
4. Injuries violence and disabilities biennial report 2004-2005. Switzerland: WHO; 2006. World Health Organization.
5. Babaei M, Mousavi S, Malek M, Tosi G, Masoumeh Z, Danaei N, et al. Cancer occurrence in Semnan Province, Iran: results of a population-based cancer registry. Asian Pac J Cancer Prev. 2005; 6 (2): 159- 64.
6. Sadjadi A, Nouraie M, Mohagheghi MA, Mousavi-Jarrahi A, Malekezadeh R, Parkin DM. Cancer occurrence in Iran in 2002, an international perspective. Asian Pac J Cancer Pre. 2005; 6(3): 359-63.
7. Burnet NG, Jefferies SJ, Benson RJ, Hunt DP, Treasure FP. Years of life lost (YLL) from cancer is an important measure of population burden – and should be considered when allocating research funds. Br J Cancer. 2005; 92:241-5.
8. Stamey JD, Young DM, Seaman Jr JW. A Bayesian approach to adjust for diagnostic misclassification between two mortality causes in Poisson regression. Statist Med. 2008; 27:2440-52.
9. Khosravi A, Taylor R, Naghavi M, Lopez AD. Mortality in the Islamic Republic of Iran, 1964-2004. Bull World Health Organ. 2007; 85:607-14.
10. Naghavi M. Death report from 29 provinces in Iran. 1st edition. 2004; Ministry of Health and Medical Education, Tehran, Iran.
11. Lyles RH. A note on estimating crude odds ratios in case–control studies with differentially misclassified exposure. Biometrics. 2002; 58:1034-6.
12. Whittemore AS, Gong G. Poisson regression with misclassified counts: application to cervical cancer mortality rates. Applied Statistics. 1991; 40:81-93.
13. Sposto R, Preston DL, Shimizu Y, Mabuchi K. The effect of diagnostic misclassification on non-cancer and cancer mortality dose–response in A-bomb survivors. Biometrics. 1992; 48:605-17.
14. Naghavi M. Death report from 18 provinces in Iran. 1st edition. Tehran, Iran: Ministry of Health and Medical Education; 2002.
15. National Statistics Center [Internet]. 2010. Available from: http://amar.sci.org.ir/PlanList.aspx
16. Pourhoseingholi MA, Faghihzadeh S, Hajizadeh E, Abadi A, Zali MR. Bayesian estimation of colorectal cancer mortality in the presence of misclassification in Iran. Asian Pac J Cancer Prev. 2009; 10:691-4.
17. Pourhoseingholi MA, Fazeli Z, Zali MR, Alavian SM. Burden of hepatocellular carcinoma in Iran; bayesian projection and trend analysis. Asian Pac J Cancer Prev. 2010; 11:859-62.
18. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin. 2009; 59:225-49.
19. Jemal A, Thun MJ, Ries LA, Howe HL, Weir HK, Center MM, et al. Annual report to the nation on the status of cancer, 1975-2005: featuring trends in lung cancer, tobacco use, and tobacco control. J Natl Cancer Inst. 2008; 100:1672-94.
20. Canadian Cancer Society's Steering Committee on Cancer Statistics. Canadian cancer statistics 2011. Toronto: Canadian Cancer Society; 2011.
21. Cancer Research UK. Lung cancer – UK mortality statistics [Internet] 2011 Jul 18. Available from: http://info.cancerresearchuk.org/cancerstats/types/lung/mortality
22. Amos A. Women and smoking. Br Med Bull. 1996; 52:74-89.
23. Mackay J, Amos A. Women and tobacco. Respirology. 2003, 8(2):123-30.
24. Brennan P, Bray I. Recent trends and future directions for lung cancer mortality in Europe. Br J Cancer. 2002; 87(1):43-8.
25. UK lung cancer mortality statistics [Internet] 2011. Available from: http://www.cancerresearchuk.org.
26. Tarrahi MJ, Mehrabani D, Khademolhosseini F, Amini M, Masoumi SJ, Julaee H, et al. Lung cancer occurrence in Southern Iran. J Res Med Sci. 2009 Mar; 14(2):139-40.